

(12) UK Patent Application (19) GB (11) 2 368 748 (13) A

(43) Date of A Publication 08.05.2002

(21) Application No 0026625.4

(22) Date of Filing 01.11.2000

(71) Applicant(s)
3COM Corporation
(Incorporated in USA - California)
5400 Bayfront Plaza, P.O.Box 58145, Santa Clara,
California 95052-8145, United States of America

(72) Inventor(s)
Alan Roger Poulter
Quang Tien Tran
Sharon Linda Wiles
David John Law
Christopher Alan Walker

(74) Agent and/or Address for Service
Bowles Horton
Felden House, Dower Mews, High Street,
BERKHAMSTED, Herts, HP4 2BL, United Kingdom

(51) INT CL⁷
H04Q 11/04 , H04L 12/56 29/06

(52) UK CL (Edition T)
H4K KTKX

(56) Documents Cited
GB 2357390 A GB 2350032 A

(58) Field of Search
UK CL (Edition S) H4K KTKX
INT CL⁷ H04L 12/56 29/06 , H04Q 11/04
ONLINE : WPI ; EPODOC ; PAJ

(54) Abstract Title
Cascade control architecture and operation for packet based communication systems

(57) A control device for a cascade connection of network units includes control logic which increments a count of active units and passes the count to the next control device. The count is initiated and terminated by an absence of control messages at the down port and up port respectively. The incremented count is passed up the cascade and a final count is passed down the cascade.

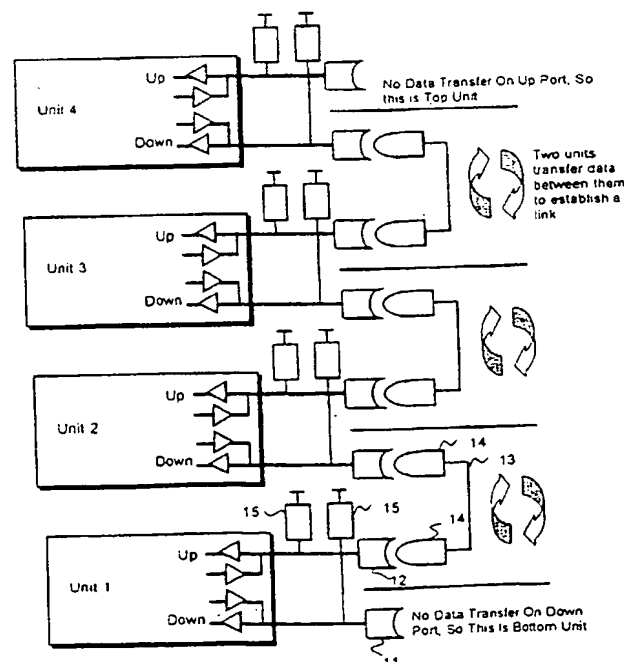


FIG.1

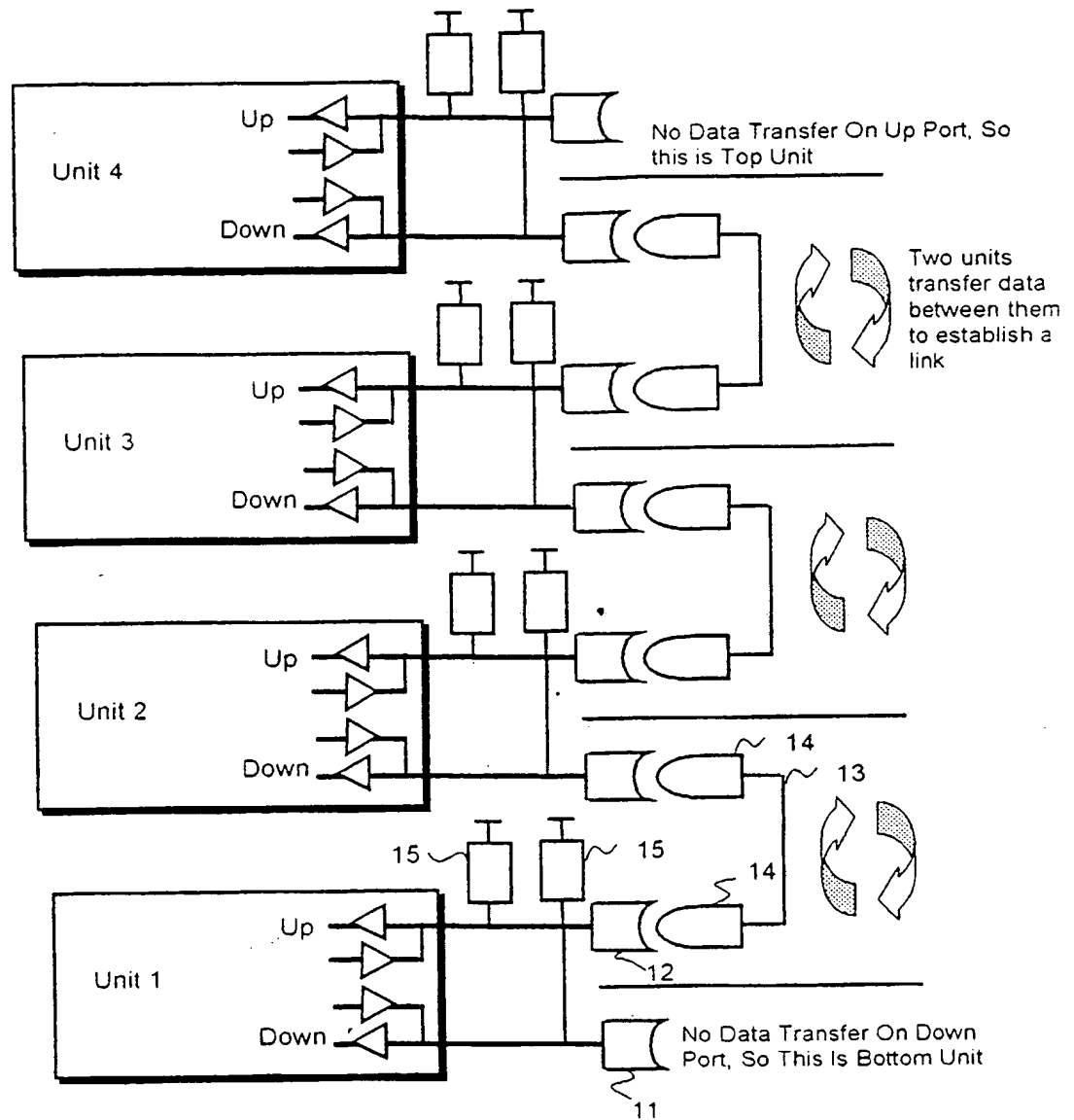


FIG.1

Down Port	Up Port	Position in Stack
No LinkOk	No LinkOk	No Cascade (Single Unit)
No LinkOk	LinkOk	Bottom of Stack
LinkOk	No LinkOk	Top of Stack
LinkOk	LinkOk	Middle of Stack

FIG.2

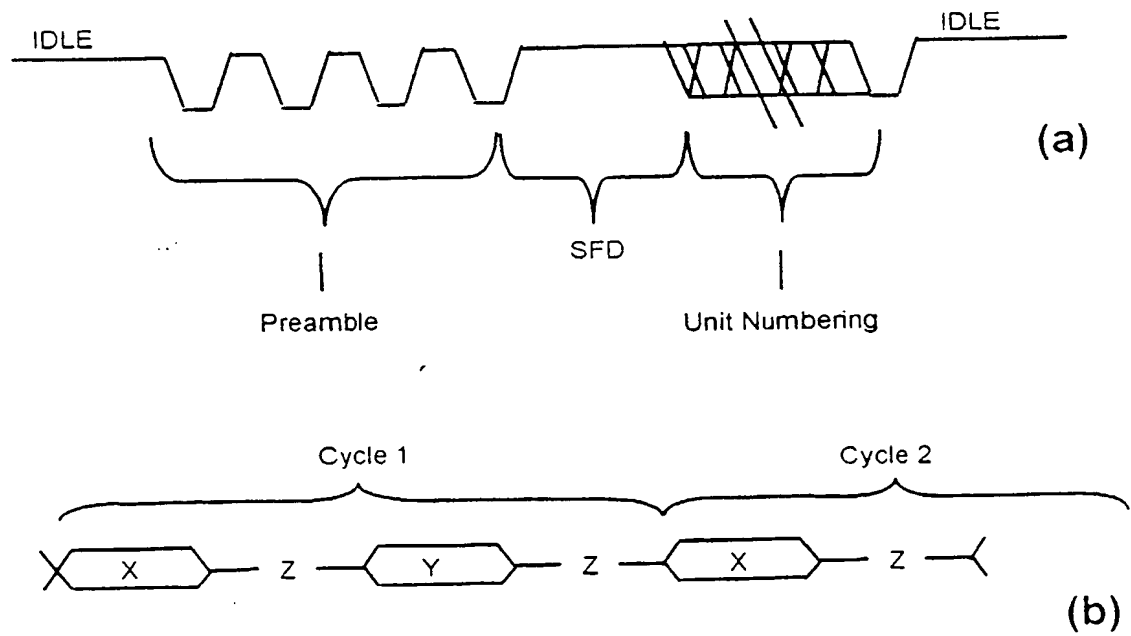


FIG.4

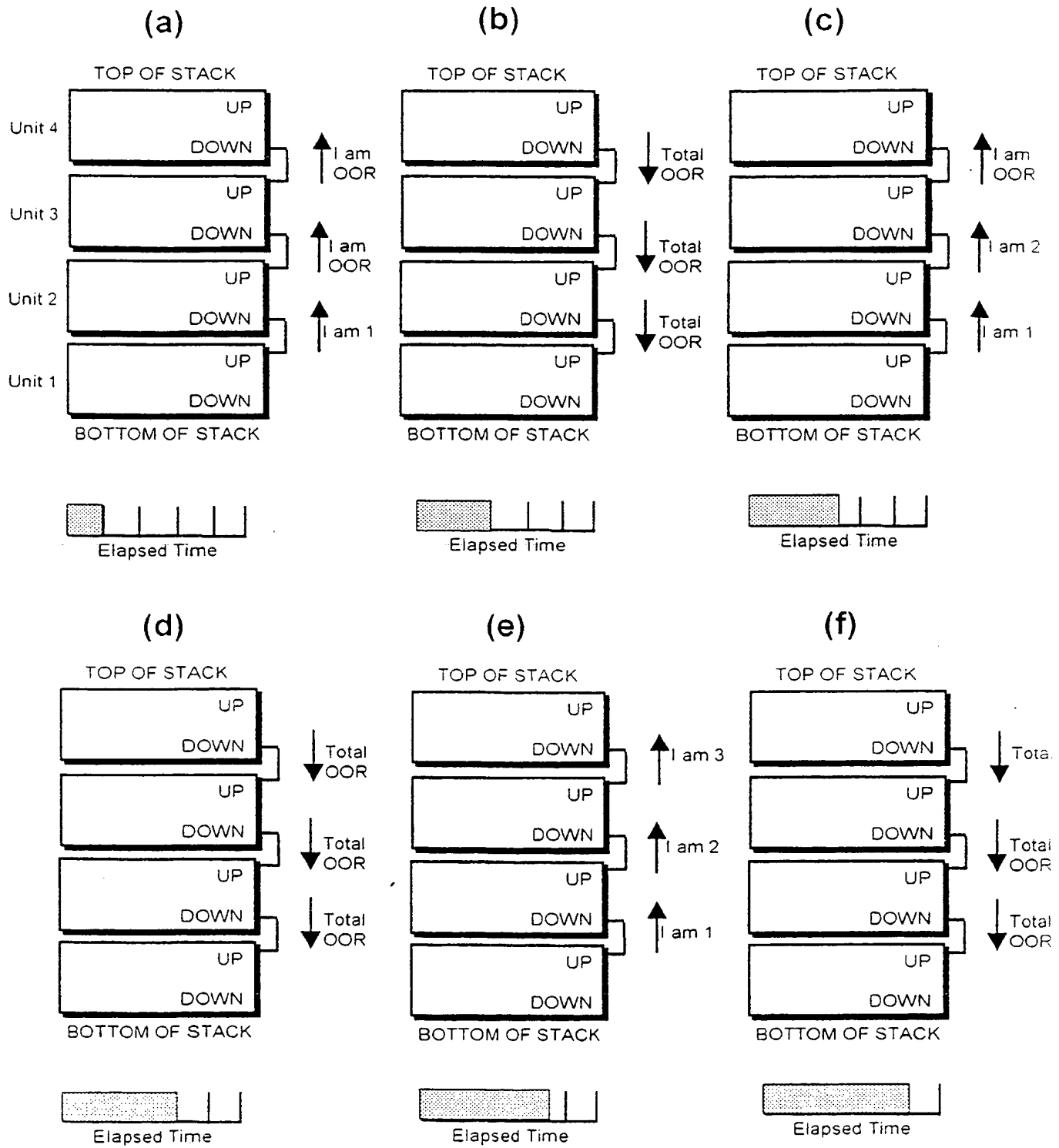


FIG.3



FIG.5

CASCADE CONTROL ARCHITECTURE AND OPERATION FOR PACKET-BASED COMMUNICATION SYSTEMS

5 This invention relates to the cascade connection of a plurality of 'stackable' units for a packet-based communication system. The invention more particularly relates to the physical architecture of the control path for such a cascade connection and the intercommunication of control messages for units which participate in the cascade.

Background to the Invention

10 The physical elements in a packet-based data communication system typically include multi-port network units such as switches and hubs. Such network units are commonly and conveniently manufactured with a fixed plurality of ports by means of which the network units are connected by way of appropriate transmission media or links to other network units
15 or users, represented by data terminals, personal computers and so on. In order to provide greater versatility in constructing the physical layout of networks by minimizing the number of different sizes of units that need to be deployed it is common to render the units 'stackable'. This term is intended to mean that a multiplicity of such units can be connected to form effectively a single controllable entity. Thus, merely by way of example, three
20 switches each with 24 ports may be 'stacked' so that effectively they constitute a single switch with 72 ports. It is customary to make such units physically stackable in a column though this physical aspect of the term 'stackable' is not essential to the meaning of the term or the present invention.

25 In order to provide intercommunication between the network units, so that for example a packet received at the port of any one of the units may be forwarded from a port or ports on a different unit or even a multiplicity of different units in the stack, depending on the addressing of the packet, it is customary to connect the stacked units together by means of a cascade, which is the term used for the communication links and associated control

functions by means of which packets are transmitted from one unit to each other unit in the stack. The cascade allows users to be connected to any of the units in the stack and needs to be organised so that the stack can be managed as a single logical entity.

5 From the operational point of view, the forwarding of packets on the cascade has to be controlled according to a variety of rules which are intended to conform to the forwarding rules relevant to different types of packets (such as unicast, multicast and broadcast) as well as other relevant rules, such as those relating to bridging. An example of such a rule is the prevention of the forwarding of a packet out on a port by which it has been received. Such
10 rules may need to be modified to make the operation of the cascade more convenient or more versatile, as exemplified in our prior Patent Applications Nos. 9929886.1 and 0010032.1.

15 Two variable values which are the basis of the operational control of a stack are the unit numbering, namely some numerical or coded quantity which identifies and distinguishes each physical unit in the stack and an 'active unit' total, which indicates the number of units actively participating in the stack. It is desirable to compute the unit numbering of the units in a stack and the active unit total and to convey the active unit total to all units in the stack automatically.

20 It is possible to employ an 'in-band' configuration in which the configuration process by means of which the units are allotted their stack identification by means of configuration packets sent during a configuration routine on the data path for packets connecting the unit in the stack. Such a system is described, for example, in GB published Patent Application
25 No. 2338155, which describes a cascade connection in which packets are employed to configure the cascade such that only one unit, called the master, is allowed to place packets on the cascade connection at any time. GB Patent Application No. 9910895 3 filed 12 May 1999 and corresponding US Patent Application for Byham et al, Serial No. 09/369, 323 describe an automatic configuration process by means of which configuration packets are

used to determine which unit is deemed to be at the bottom of the stack notwithstanding the provision of resilience or possible misconnection of connecting cables.

5 Our co-pending Application No. 0018713.8 filed 1 August 2000 and entitled 'Cascade Architecture and Operation for Packet-based Communication Systems' describes a system, devised by various inventors common to the inventors of the present invention, in which control messages travel around the cascade connection by a control path distinct from the path of data packets and are employed in the establishment of the identification numbers for the units in the stack and are also employed for computing an 'active unit total'. The system
10 described in that application is a particularly versatile system which can accommodate the 'hot' insertion or removal of active units and which can be employed to compute the 'active unit total' which takes into account the units which are within the cascade but which may be powered down or otherwise not actively participating in the cascade.

15 The physical architecture required in that co-pending application is, as far as its preferred implementation is concerned, based on the use of special three-port connectors which define paths for the packets and the control messages and is furthermore organised so that if a unit which is connected to a connector is powered down or inoperative, the system still provides, by means of the relevant connector, both an unbroken data path and an unbroken control
20 path around the cascade. The earlier application also describes 'cascade modules' by which a network unit can be connected in the cascade.

There is however a need for a system which provides at least some of the functionality of the earlier system, and in particular the computation of unit numbering and an active unit
25 total, and which can be implemented in circumstances where the earlier system is either not feasible or not needed. For example if the available terminal connectors have insufficient terminals for supporting the variety of signals and control paths required in practice for the system described in the co-pending application, the earlier system cannot be satisfactorily employed. Furthermore, the 'hot swap' facility might not be needed. As will become more

apparent, the present invention may be implemented in a simple manner without a 'hot swap' facility, though that facility may be provided within the scope of the invention.

Summary of the Invention

5 The present invention is based on the provision of a control path for the units in the cascade, the control path being employed for the conveyance of control messages that represented an accumulated stack identification number and active unit total. However, the architecture is organised so that the absence of valid control messages at particular ports is
10 employed to determine whether the relevant network unit is the first (that is to say the 'bottom') unit in a stack or the last (that is to say the 'top') unit in a stack. This information can be used to start and terminate the process of the stack identification numbering and the computation of an active unit total. The invention may be implemented within stackable units such as switches but could also be provided in modules for use with such units.

15 Other and more particular features of the invention will become apparent from the following detail description with reference to the drawings.

Brief Description of the Drawings

20 Figure 1 illustrates a stack of communication units connected by a cascade according to the invention.

Figure 2 is a table showing the interpretation of control messages by the units in the stack.

25 Figure 3 represents the stack at various stages of a configuration process.

Figures 4(a) and 4(b) illustrate the physical nature of control messages preferably employed in the preferred example.

Figure 5 illustrates a control device according to the invention, arranged to generate and respond to control messages.

5 **Detailed Description of a Preferred Example**

Figure 1 illustrates a stack of four multi port communication units which may each comprise a multiplicity of switches, similar for example to Figure 1 of the aforementioned co-pending application. Typically, each switch may be a multi port switch which has a multiplicity of
10 ordinary ports for connection to users and, this example, 'cascade' ports by means of which the unit is connected so as to supply packets to or receive packets from other units in the stack.

For convenience, the units in the stack are denoted 'Unit 1', 'Unit 2' etc. Also, for the sake
15 of simplicity the Figures illustrates a system in which the hardware components (and particularly the control device in Figure 5) required for the invention are located within each of the units.

Each of the units has an 'up' port and a 'down' port. Within each unit the up port and the
20 down port are connected to logic circuits which perform appropriate processing of packets which are to be forwarded from the unit onto the cascade or are received from another unit by way of the cascade. Although such processing is essential to a practical system, the manner in which the processing is organised and other things such as the imposition of forwarding rules and so on are not directly relevant to the present invention.

25 Each of the down and up ports in each unit is provided with a multi-pin terminal. These are shown as terminals 11 and 12 respectively for the bottom Unit 1. The up port of a unit is connected to the down port of the next by way of a cable 13 in known form terminated at both ends by means of a connector 14 which is compatible with the terminal on each unit.

Such a cable needs to be able to convey packets in the relevant format on respective lines and needs also to have at least one line for the connection of control messages. The present invention in its preferred form is intended to be implemented using only one control line which will therefore require only a minimum of pins on the port connectors.

5

A cascade connection is made by connecting the up port of each unit to the down port of the next in turn, leaving the down port of one unit, in this example the down port of the Unit 1, and the up port of another unit, in this example the up port of Unit 4, unconnected.

10

For the sake of simplicity, only the control line, namely the single control line 13 used for the control messages to be described, is shown for each connector in Figure 1. This presumes that the units have a common datum (such as earth) for control signals conveyed on the line.

15

Shown in Figure 1 and intended for further explanation below is a clamping (e.g. pull up) resistor 15 for each control line to a VCC bus within each of the communication units, Unit 1 to Unit 4.

20

Figure 1 also shows the basic operational scheme of the invention. It will be seen that the up port of Unit 1 is connected to the down port of Unit 2, the up port of Unit 2 is connected to the down port of Unit 3 and the up port of Unit 3 is connected to the down port of Unit 4. Thus there is a control path by means of which, at different times, the respective adjacent units can exchange control messages. However, the down port of Unit 1 and the up port of Unit 4 are not connected and are therefore unable to receive valid control messages. This fact is exploited by logic within the units to establish Unit 1 as the bottom unit in the stack and Unit 4 as the top unit in the stack.

25

Preferably, those units which have a connection from the up port of one to the down port of the other exchange control messages so that the logic within the units establishes whether

the link between the respective units has been established or not. The relevant states are 'LinkOk' and 'no LinkOk' respectively. Figure 2 summarises the status of the ports and it will be seen that if 'LinkOk' is established for both ports of a unit the unit is deemed to be in the middle of a stack whereas if there is no LinkOk signal developed for the up port and the down port then the unit is deemed not to be part of a cascade.

Figure 3 illustrates the stack, shown in this particular example as a stack of four, at various stages in the preferred configuration process. Each of the units will at relevant times transmit a signal that contains a stack identification number. In this example the stack identification number will correspond to the active unit count and the stack identification number of the top unit will eventually correspond to the active unit total.

Figure 3 (a) shows at the end of a first cycle the transmission of a signal from the up port of Unit 1 to the down port of Unit 2. Unit 1 has already determined, by virtue of the presence of LinkOk status for its up port and no LinkOk status for its down port that it is at the bottom of the stack and accordingly it can transmit a signal which for the sake of simplicity is represented as 'I am 1' from its up port to the down port of Unit 2. At this time Unit 2 does not yet know the total number of units and accordingly by default it transmits an 'out of range' number, shown herein as 'OOR'. Thus from its up port it transmits a signal which is called for simplicity 'I am OOR'. Unit 3 transmits a similar signal from its up port to Unit 4.

In practice, as will be seen later, a message such as 'I am X' is constituted by a numerical value (X) in a dedicated field in a formatted control message.

In the next cycle, since the number of units has still not been established, the units transmit from their down ports to the up port of the next unit below a signal which is shown as 'Total OOR', representing an out-of-range (invalid) value for the active unit total.

Figure 3 (c) shows the next phase in which Unit 1 continues to transmit 'I am 1' from its up port to the down port of Unit 2 whereas now Unit 2 transmits 'I am 2' from its up port to the down port of Unit 3. Unit 3 still transmitting 'I am OOR' from its up port and down port of Unit 4.

5

Figure 3 (d) shows the next phase in which, since the total of units that has not yet been established, there is a transmission of a 'Total OOR' from the top of the stack downwards.

Figure 3 (e) shows the next phase in the configuration cycle wherein in addition to the signals transmitted from Unit 1 and 2 in Figure 3 (c), Unit 3 is transmitting 'I am 3' from its up port to the down port of the stack.

10

Finally, Figure 3 (f) shows the transmission of a 'Total 4' from the unit at the top of the stack downwards.

15

This last figure shows 'Total OOR' transmitted from the down ports of Unit 3 and 2. During following cycles the message 'Total 4' will be transmitted down the rest of the stack until the up ports of Units 2 and 1 receive the message 'Total 4'.

Figure 4 (a) shows a control message of the type which may be employed in the present invention. This control message resembles that shown in the aforementioned application. It is in the form of a packet comprising a preamble, typically four bits at a bit rate different from the rest of the numbering, a start frame delimiter, which may be a short period at a tri-state value, followed by several data bytes transmitting either up port data or the down port data. The 'up port' data is a number identifying the respective unit. The 'down port' data is a number which represents eventually an 'active unit total' though, as already explained the total will initially be 'out of range'.

20

25

Other fields could be included in the control messages broadly on the lines as described in the co-pending application of previous mention.

5 The logic in each unit therefore needs to generate, in the format shown for example in Figure 4, the messages shown in Figure 3.

10 If the unit is at the bottom of the stack, the up port data will signify the message 'I am 1' and therefore the 'X' data represents '1' which is the active unit count at the up port of Unit 1 and also the numbering of Unit 1. The logic will (if the unit is not at the bottom of the stack) produce up port data representing an out of range number (OOR) until the unit's logic receives an in-range number at its down port and will then increment that number by unity.

15 The logic is also required to produce down port data (Y). If the unit is at the top of the stack the 'down port data' is initially 'out of range' and eventually the active unit total. If the unit is not at the top of the stack the down port data corresponds to the data which the unit receives at its up port, as described with reference to Figure 3.

20 Figure 4(b) illustrates a preferred cycle of generation of the control messages (packets). In each cycle, a control packet containing up port data (X) is followed by a high impedance (Z) state, and then a control packet containing down port data (Y) is followed by the high impedance (Z) state.

25 Figure 5 illustrates the control device which in this example may be provided for each of the network units in order to generate and translate the control messages.

The circuits shown in Figure 5 has an input/output down port 51. On the input side of this terminal there is a connection by way of a pull up resistor 52 to a positive rail VCC. The port is also connected to one input of a gate 53 which can receive enable signal on a line 54

extending also to an enable input of a tristate buffer 55. This enable signal is obtained from a packet transmit circuit (PKT TX) 56.

5 The output of OR gate 53 is connected to a packet receive (PKT RX) circuit 57 which is coupled to one input of a multiplexer 58, the signal from the circuit 57 being coupled to the output of the multiplexer if the packet receive circuit produces a LinkOk enabling signal.

10 The output of the multiplexer is coupled to a 'unit ID' register 59 the output of which is incremented by one in an incrementer 60 and coupled to a packet transmit (PKT TX) circuit 61. One output of this packet transmit circuit goes to the input of a tristate buffer 62 which is connected to an input/output up terminal 63 provided with a pull up resistor 68 to the positive rail.

15 The circuit in the reverse direction is somewhat similar. The input/output of the up terminal is connected to an OR gate 64 of which the other input is provided with an enable signal from the packet transmit circuit 61. The output of the gate 64 goes to a packet receive (PKT RX) circuit 65 of which the output is coupled to an input of a multiplexer 66. The multiplexer can provide a LinkOk signal to the multiplexer 66, which drives a register 67.

20 The blocks (57,65) labelled PKT RX perform the decoding of the control messages. Their input is the serial management packets and they can provide three outputs: a linkOK output which indicates continuing reception of the serial management packets, a pktRcvd output which produces a pulse or toggle signal when a control packet is received and a data value which represents the data field extracted from the received packet. If the control packet is
25 received at an up port (63) the data field will be the active unit count value from the above unit; if the control packet is received at the down port (51) that data field will be the unitID value from the unit below. The 'pktRcvd' signal is not used on the up port.

The blocks (61,56) labelled PKT TX perform the encoding of the control messages. They have two inputs: a data value and a transmit control signal. For an up port, the data value is the unitID value to be transmitted up to the above unit in the stack. For a down port the data value is the active unit count value to be propagated down the stack. The transmit control signal forces the sending of a control packet (typically a pulse or toggle signal would drive this line). For the up port this signal is driven at regular intervals (by the txPkt signal which would be directly connected to some kind of timer circuitry), thus causing the sending of control packets at regular intervals. For the down port the transmit control signal is driven by the pktRcvd signal from the down PKT RX block. Thus, control packets continually sent on the up port but are only sent on the down port in response to control packets being received on the down port, i.e. the down port is the slave and the up port is the master.

The outputs of each PKT TX block are the serial control packets (containing the supplied data value) and a transmit envelope signal (txEnv). This 'txEnv' signal is active high while packet transmission is taking place and it enables the data onto the bidirectional serial data line. Between packet transmissions 'txEnv' is low and the serial data line is not driven so it can then be used for packet reception. Also, during transmission when 'txEnv' is high, packet reception is blocked by ORing the received data with txEnv (OR gates 64 and 53) - this ensures that a port does not receive its own transmissions, because (owing to the pull-up resistor) the logical 1 state is the idle state between packets.

The operation of the multiplexers is as follows:

If there is no signal received on the down port (i.e. unit is bottom of stack) PKT RX (57) produces no 'linkOK' signal and multiplexer 58 forces a value of zero into the unit ID register 59 (unit 0 is bottom of stack). UnitID is incremented by incrementer 60 before being sent up to the next unit in the stack. This unit, seeing linkOK, loads the received unitID value (unit 1) into its unitID register.

A unit at the top of the stack will not see any 'linkOK' signal in respect of its up port. Thus, the multiplexer 66 will route the transmitted unitID value into the active unit count register 67. This is correct: the total number of units in the stack is equal to the unit ID of the top unit plus 1 (assuming that numbering starts from zero). All units other than the top unit will see a 'linkOK' in respect of their up port and each will copy the active unit total value from the unit above into its active unit total register 67. Accordingly this value for the active unit total will propagate down the stack.

The unit ID register 59 and active unit count register 67 can be reset on power-up to an out of range value (e.g. for a 4 bit counter set to all ones). The increment function (block 60) will not increment this value and the system will recognise it as being out of range, and thus invalid if it is read as a unitID or active unit count.

In the foregoing the control device shown in Figure 5 has been described as incorporated within the network units (e.g. switches) for which the cascade connection is formed. This form of the invention does not provide a 'hot swap' facility. However, by incorporating the control devices in respective modules separable from the network units and making minor consequential modification it is feasible to retain a 'hot swap' facility. Such modules may be in the general form described in the aforementioned copending application and may provide in addition to the logic described in Figure 5 transmission paths for data packets to and from the cascade.

Claims

1. A control device for a cascade connection of a multiplicity of network units, the control device including an up port and a down port and control logic for sending and receiving control messages to and from the up port and the down port, the control logic including:
- 5 means for incrementing a count value in a control message received at the down port to provide for the up port a control message including an incremented count;
- 10 means responsive to a control message including a count value received at the up port to provide a corresponding control message for the down port;
- means responsive to the absence of reception of valid control messages at the down port to initiate a count for a control message at the up port; and
- 15 means responsive to the absence of reception of valid control messages at the up port to determine a total count value in a control message for the down port.
2. A control device according to claim 1 and arranged to receive and transmit control messages on the same line from each of the up port and down port.
- 20 3. A control device according to claim 1 or 2 wherein the control logic provides for the up port a control message defining an invalid count until a valid count is detected in a control message received at the down port.
- 25 4. A control device according to claim 3 wherein the control logic provides for the down port, in the absence of reception of valid control messages at the up port, and in the absence of a valid count in a control message at the down port, a control message defining an invalid count total.

5. A stack of multi-port network units including a cascade connection for the conveyance of data packets from each of the network units to the other network units in the stack and for each network unit a respective control device according to any foregoing claim, disposed to establish for the network units respective identification numbers corresponding to successive values of the count and an active unit total corresponding to said total count value



INVESTOR IN PEOPLE

Application No: GB 0026625.4
Claims searched: 1-5

15

Examiner: Richard Howe
Date of search: 22 August 2001

Patents Act 1977 Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:
UK Cl (Ed.S): H4K (KTKX)
Int Cl (Ed.7): H04L (12/56 , 29/06) ; H04Q (11/04)
Other: Online : wpi ; epodoc ; paj

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	GB 2 357 390 A (3Com) - see abstract	
A	GB 2 350 032 A (3Com) - see abstract	

N	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.